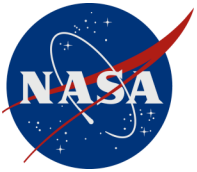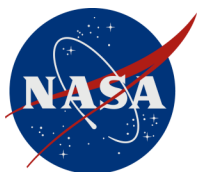# NASA's Earth Science Data Life Cycle

Ken McDonald
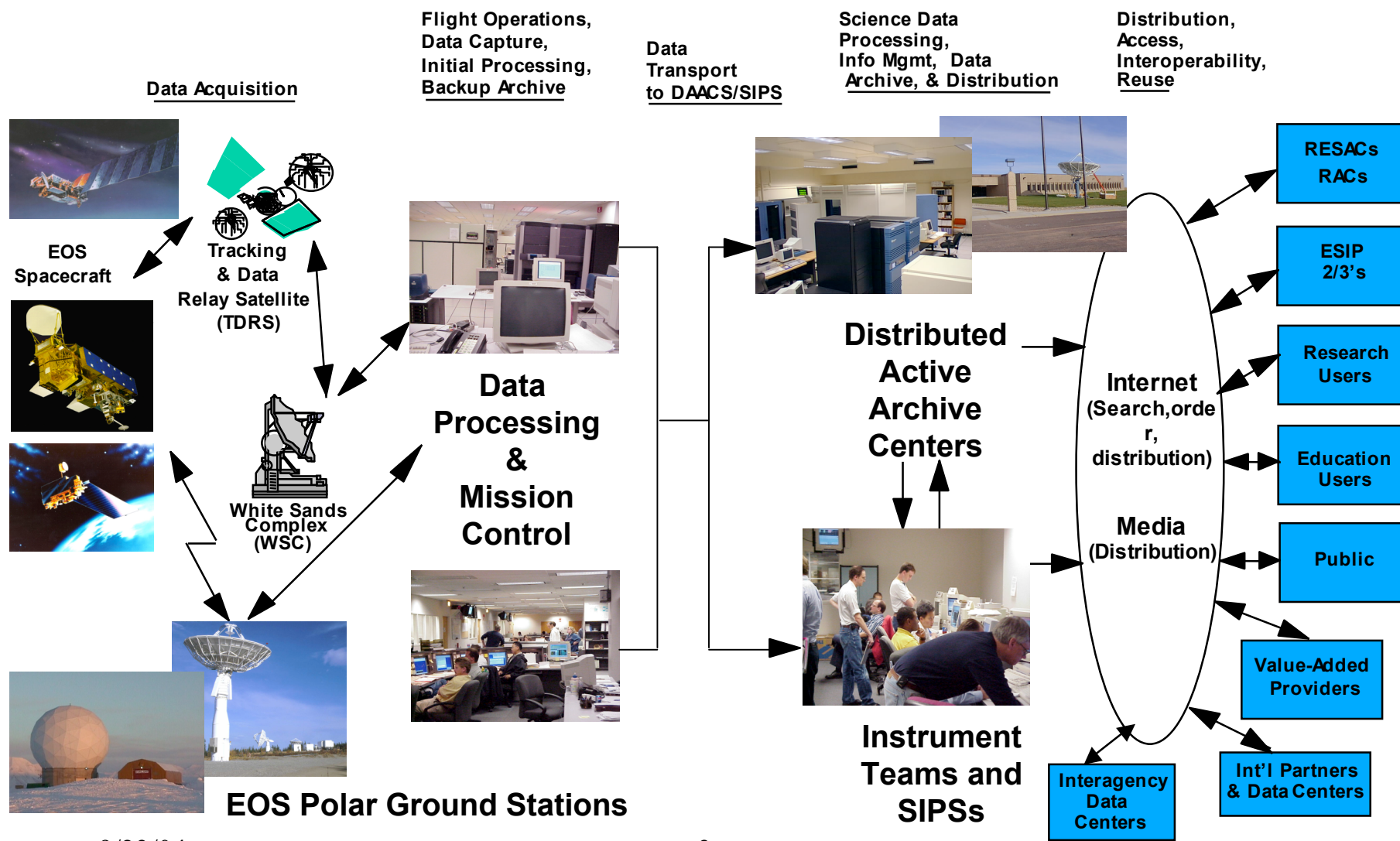# Earth Science Data and Information System Project
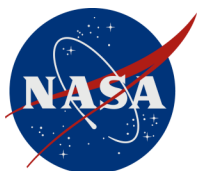
Background - Requirements

- NASA shares the responsibility for stewardship of its Earth science data resources with NOAA and USGS.
  - NASA holds the responsibility for its data during the life of each mission plus four years.
  - NOAA and USGS provide the long term archive for ocean and atmosphere data and land processes data, respectively.
- Close coordination and cooperation among the three agencies is required to ensure the long-term preservation and utility of the data.
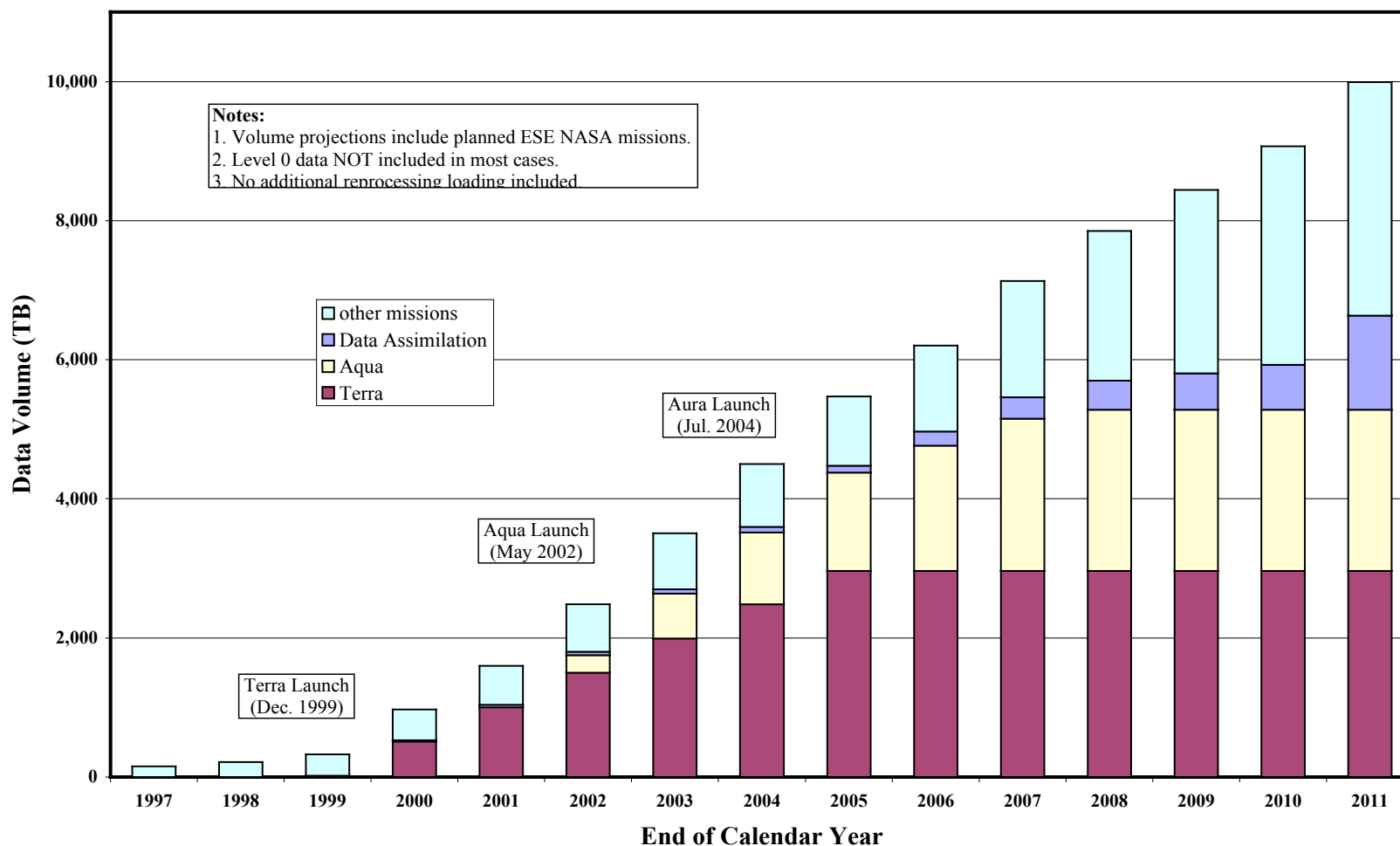
NASA Responsibilities

Data Acquisition

Flight Operations, Data Capture, Initial Processing, Backup Archive

Data Transport to DAACS/SIPS

Science Data Processing, Info Mgmt, Data Archive, & Distribution

Distribution, Access, Interoperability, Reuse

EOS Spacecraft

Tracking & Data Relay Satellite (TDRS)

White Sands Complex (WSC)

Data Processing & Mission Control

EOS Polar Ground Stations

Distributed Active Archive Centers

Instrument Teams and SIPSs

Internet (Search, order, distribution)

Media (Distribution)

RESACs RACs

ESIP 2/3's

Research Users

Education Users

Public

Value-Added Providers

Interagency Data Centers

Int'l Partners & Data Centers

NASA Earth Science Data Holdings

Total Accumulated Product Volumes from ESE NASA Missions (1997-2011)

Notes:
1. Volume projections include planned ESE NASA missions.
2. Level 0 data NOT included in most cases.
3. No additional reprocessing loading included.

Legend:
- other missions
- Data Assimilation
- Aqua
- Terra

Aura Launch (Jul. 2004)

Aqua Launch (May 2002)

Terra Launch (Dec. 1999)

Y-axis: Data Volume (TB)
X-axis: End of Calendar Year (1997–2011)
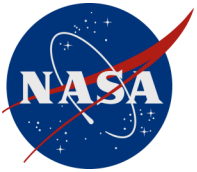
Past Activities and Milestones

- ESDIS, DAACs and Instrument Teams supporting active archive functions.
  - Production of standardized products and metadata.
  - Generation of data set documentation.
  - Archive and documentation of data production software.
- NASA, NOAA and USGS have signed interagency agreements in place for long-term archive transfers.
- Several community workshops have addressed data lifecycle issues.
- Data lifecycle was one of the study areas of the Strategic Evolution of ESE Data Systems (SEEDS).

Data Lifecycle - Community Input

- ## Multiple workshops and studies.
    - "Global Change Science Requirements for Long-Term Archiving", Report of the Workshop, Oct. 28-30, 1998, U.S. Global Change Research Program (USGCRP) Office, March 1999.
    - "Ensuring the Climate Record from the NPP and NPOESS Meteorological Satellites", National Research Council Committee on Earth Studies (NAS-CES) 2000.
    - "Workshop on Long-Term Archiving of EOS Data", Jan. 29-30, 2002, EOS Science Working Group on Data (SWGD).

- ## Data Lifecycle Study and community workshops of the "Strategic Evolution of ESE Data Systems (SEEDS)" formulation activities.
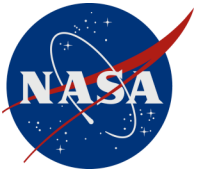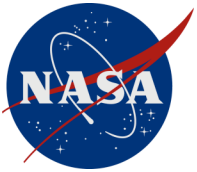
NASA/NOAA LTA Agreements

Within the 1989 NASA-NOAA MOU NASA has agreed to:

Use its best efforts to ….. Transfer to NOAA, at a time to be determined, responsibility for active long-term archiving and appropriate science support activities for atmosphere and oceans data, as defined in accordance with approved coordinated program definition, development, and implementation activities and plans.

And conversely, the NOAA has agreed to:

Use its best efforts to ….. Assume responsibility at a time to be agreed upon for active long-term archiving and appropriate science support activities for atmospheric and oceans data for the EOS program.

NASA/USGS LTA Agreements

Within the 1988 NASA-USGS MOU NASA has agreed to:

Transfer responsibility for active long-term archiving and appropriate science support activities for these data archives to the USGS at agreed upon future dates in accordance with approved coordinated program definition and implementation plans.

And conversely, the USGS has agreed to:

Assume responsibility at a future agreed upon time for active long-term archiving and appropriate science support activities for those NASA experimental land remotely sensed data in the active short-term archives in accordance with approved coordinated program definition and implementation plans.

SEEDS Data Lifecycle Study

- A NASA/USGS team conducted the study with input from NOAA and the ESE data provider and user communities.

- A Data Life Cycle section of the SEEDS Draft Recommendations document has been published.
  - Identifies purpose, goals and objectives of DLC efforts.
  - Proposes "responsibilities" for different facilities that handle data and information as it migrates over its life cycle.
  - Describes preliminary set of issues and next steps.

- Study strongly recommends the formation of a Data Lifecycle Working Group to guide transitions of data.
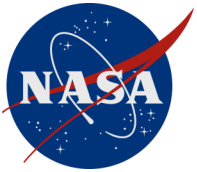
Agreements vs. Actions

- At a high level, responsibilities and requirements have been addressed.
  - MOUs, agency plans and schedules, community needs, etc.
- Work on meeting those responsibilities and agreements is just beginning.
  - Activities underway with both USGS and NOAA.
  - Approaches are different because situations are different.
  - In both cases, implementation is more bottom up, governed by high level goals and schedules.
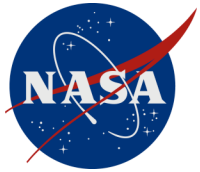
Current Status - NASA/USGS

- Version 0 datasets transferred to USGS responsibility
    - Advanced Solid-state Array Spectrometer (ASAS)
    - Advanced Very High-Resolution Radiometer (AVHRR)
    - Global Composites
    - Normalized Difference Vegetation Index (NDVI) CD
    - Orbital Segments
    - Aircraft Scanners
    - Global Land Cover Characterization (GLCC)
    - Global Land Cover Test Sites (GLCTS)
    - GTOPO30 and Hydro 1k
    - NASA Landsat Data Collection (NLDC) MSS/TM
    - North American Landscape Characterization (NALC)
   - Spaceborne Imaging Radar-C (SIR-C) and Educational CD

Current Status - NASA/USGS

- Landsat-7 operations and funding responsibility taken over by USGS

- EOS missions – data are at EDC; transition is simple
  - Advisory committee's prioritization
  - Funding

- Access will be through USGS's system (http://edc.usgs.gov/)

Current Status - NASA/NOAA

- Funding has been an issue for several years
  - NASA has worked with NOAA to help define funding requirements and funded small prototyping efforts.
  - The Comprehensive Large Array-data Stewardship System (CLASS) defined to support NOAA and NASA archive needs.
  - NOAA obtained CLASS funding in 2001
  - FY 2004 is first year NOAA has had funding for addressing EOS data
- Plans to initiate transition of MODIS data being developed
- NASA/NOAA joint working group meeting periodically to work out interfaces
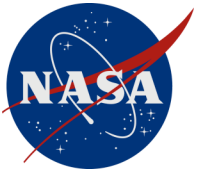
CLASS Overview

CLASS is a web-based data archive and distribution system for <u>ALL</u> NOAA/NESDIS environmental data

<u>**Archive**</u> … ingest, storage, metadata management, and data quality assurance

<u>**Distribution**</u> … access, visualization, and data delivery

CLASS currently supports POES, DMSP, and GOES data sets, plus RadarSat (Synthetic Aperture Radar) and SeaWiFS (Ocean Color Product)

CLASS will support additional campaigns, broader user base, new functionality currently being defined

- EOS is one of the additional campaigns, along with NPP, NPOESS…

Richard G. Reynolds (OSD)

Charles Bryant (OSD), Chung Wu (OSD)

Suitland Design & Development Team

Alex Kidd (OSD)
Robert Rank (OSD)
Tino Cremidis (CSC)
David Bowman (NCDC)

- Baseline system
- Processes/policies
System Architecture
& Design
- System Integration
& Testing
- Support Suitland
operations

West Virginia Development Team

Carlos Martinez (TMC)

- GOES Lead
- CCR Implementation
- NPP Implementation
- EOS Implementation
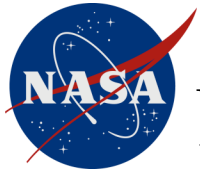- Support Asheville
operations

Boulder Development Team

Eric Kihn (NGDC)

- SABR
Integration
- DMSP
- Data Mining

•Ted Habermann (NGDC)

- Geospatial
databases
- Geotiff images
- Metadata

NVDS & COAST Leveraging Team

Chung Wu (OSD)

- E-commerce
- Order
Management
System

NOAA/NASA CLASS WORKING GROUP

Formed in FY 03 by mutual agreement.

NOAA Suitland team led by Bob Rank

- Also includes representatives from Fairmont WVa and NGDC in Boulder, CO

NASA team led by ESDIS / Ken McDonald

- Includes additional ESDIS representatives, Goddard DAAC and SWGD Chair
- HQ, Martha Maiden et al frequently sits in on meetings.

Purpose of group:

- Define the interface between EOSDIS and CLASS.
- Define an initial set of requirements.
- Plan and begin the first series of transfers.

Group met in spring/summer '03 but put itself on hold when NOAA lost funding dedicated to EOS campaign.

Working Group Activities

Established MODIS L0 and L1B as initial candidates for data transfer.

- L0 is stable and L1B has high user demand.

NASA and NOAA network engineers have been added to the Working Group.

NASA has action to generate "Data Submission Agreement"

- Outlines POCs, data and metadata descriptions, data volumes, data services, etc.

NOAA has action to develop a schedule to test the interface and initiate data transfers.

- Tentative plan is to take advantage of the MODIS Collection 5 reprocessing campaign that begins in November and runs for 12 months (will probably begin some time during the campaign).
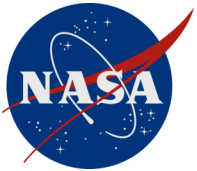
Issues

- Representation and interrelationships of panels and committees
  - NASA, NOAA, USGS Steering Committees
- Coordination of community expectations with interagency responsibilities
- Effect of transitioning to measurement-based systems on data life-cycle
  - No longer a mission based schedule.
  - Reprocessing to create long time-series.
- Preservation of software and its provision to users upon request
  - Open policy versus intellectual property rights.
- Incentives for supporting the data transfer.